Sociology 2275, Social Network Analysis, Fall 2022

Assignment 3: Picturing Networks

This assignment gives you a bit of experience in working with sna visualization software in R, and in using pictorial displays to explore how location within a network reflects different actor attributes.

For this assignment, work with a small data set about an organization in the entertainment industry known as "The Corporation." I learned about these data from

> Lusher, Dean and Garry Robins. "Example: Exponential Random Graph Model Analysis." Pp. 37-46 in Lusher, Dean, Johan Koskinen and Garry Robins, eds. (2013). *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*. New York: Cambridge University Press.

Nodes are the 38 executives in the study. Arcs with a value of "1" indicate that the row executive cited the column executive as someone with whom it was important to communicate in order to accomplish work, while a 0 indicates that no such citation was made.

The data project also includes attribute data about the 38 executives. The 3 attributes are, respectively, as follows:

> (1) office, categorical indicator of the branch office in which an executive was based, varies from 1-3;
> (2) projects, number of projects completed (this can be taken as an indicator of experience);
> (3) seniority, a binary seniority indicator, presumably with 1 denoting more senior and 0 less senior executives.

The data are stored in a data project called Corporation that is posted in the Data tab of the Modules folder of the course Canvas site. It is stored as an .RData file; you should begin by loading it into R ("`load`" command in R, or "`File>Open File`" in RStudio, and saving the data in a location that you can remember within your R workspace.

A support handout in the Implementation/Software notes tab of the Modules folder on the course Canvas site offers some guidance about how to use the graphics function `gplot()` in package sna; you should rely on that for this assignment.

Bold points below indicate what you should turn in.

1. **Calculate and report a subgroup density table** for the directed communication network, partitioned by seniority. **Do communication relationships tend to occur among executives who have similar seniority?**

2. Calculate the indegree for each actor and save the results in a file or vector (you can do this via the `xDegreeCentrality()` function in xUCINET. **Which executive receives the most citations as an essential communication contact? Which receives the fewest?**

3. Using the `gplot()` function, obtain the default (Fruchterman-Reingold) spring-embedder representation of the communication network. Then "decorate" the plot by

a. letting the color of the vertices represent seniority differences, and
b. letting the shape of vertices correspond to branch office.

Turn in a copy of the decorated plot. Do more senior executives appear to be more central within the communication network? Which *less senior executives* are most often regarded as essential communication contacts?

You *may* wish to experiment by further decorating the plot, say letting the size of nodes correspond to indegree; you would use the vector of indegrees to govern vertex size. (This activity is optional, nothing need be turned in.)

4. Use the attribute data to investigate which attribute (office, projects completed, or seniority) appears most closely related to the communication pattern within the firm. **Turn in a copy of a decorated plot that you find most informative for understanding the communication network in the firm, and summarize its content in one paragraph.**

Due: Tuesday, September 27.

Note: These assignments are not graded, but we do take note of whether or not you do them. Please submit to either derick_baum@g.harvard.edu or pvm@wjh.harvard.edu .